Form O. 38.
5 rms. S.O.S. 2/35. S4708

The Incorporation of National Institutions for Persons requiring Care and Control.

LONDON Offices of the Incorporation:
14, HOWICK PLACE, WESTMINSTER, S.W.1
Telegrams: Burdensome, Sowest. London.
Telephone: Victoria 3045.

Colony Telephone: FISHPONDS 54.

1.10.42

I beg to certify as follows:

First, that in regard to the manuscript work entitled "Blood-Group Frequencies in South-Western England and North Wales: A Study in Racial Variation, together with a Search for Evidence that the Blood-Groups Possess Selective Value", submitted to the University of Edinburgh as part of a Thesis for the degree of Doctor of Medicine, the material, in the form of records of the grouping of blood-donors, was made available by the authorities acknowledged. The methods of tabulation were devised by myself, and the work carried out by me with the help of my assistant, Mrs. Lenton. The calculations were performed and the manuscript written by myself.

Secondly, that the authorship of the published works submitted is indicated by their titles. In the case of publications bearing my name alone the work was carried out and publications written by myself. Acknowledgements for any assistance has been made in each case. The work described in the series of five papers - Studies on a Child Population - was planned by me and carried out with the assistance of my colleagues whose names appear in four instances. The papers were written by me with suggestions from them. In the case of the papers with Prof. C. Bruce Perry and with Dr. R.G. Gordon and Dr. Ruth Griffiths my share was the statistical analysis and the description of the findings from that point of view.

BLOOD-GROUP FREQUENCIES IN SOUTH-WESTERN ENGLAND AND NORTH WALES: A STUDY IN RACIAL VARIATION, TOGETHER WITH A SEARCH FOR EVIDENCE THAT THE BLOOD-GROUPS POSSESS SELECTIVE VALUE.

Submitted to the University of Edinburgh, together with previously published works, as a Thesis for the degree of Doctor of Medicine.

by

John Alexander Fraser Roberts.
M.A.,M.B.,CH.B., D.Sc.,F.R.S.E.

Director, Burden Mental Research Department,
Stoke Park Colony, Bristol.

Special Lecturer on Human Genetics, University of Bristol.

Probationary Temporary Surgeon-Lieutenant, R.N.V.R.

October, 1942.

CONTENTS.

I. INTRODUCTION.

It was first discovered by L. and H. Hirszfeld (1919) that the races of mankind differ in the relative frequencies of the four classical blood groups. Since that time an enormous amount of information has been collected from all parts of the world. To-day it can be said that more is known about the geographical variations of the human blood group genes than is known in the case of any other genes whatsoever, whether plant or animal (Dobzhansky, 1941).

The four original blood groups depend upon the presence or absence of two agglutinable substances, or agglutinogens, in the red blood corpuscles, associated with the presence or absence of corresponding agglutinins in the serum. The agglutinogens are usually denoted by the letters A and B, the agglutinins by the letters a and b. Red cells containing A are agglutinated by serum containing a; similarly, red cells containing B are agglutinated by serum containing b. There are four possible combinations:

Group	Cells	Serum	(Old Moss classification.)
AB	Both A and B.	Neither a nor b.	(I)
A	A	b	(II)
B	B	a	(III)
O	Neither A nor B.	Both a and b.	(IV)

Persons of group O are universal donors, for their red cells do not contain either agglutinogen and so cannot be agglutinated by the serum of any recipient.

The inheritance of the blood groups is simple. The mode of transmission was first elucidated by Bernstein (1924, 1925), and his hypothesis is now universally accepted. Three genes, A, B and O, form a series of triple allelomorphs, situated at a particular locus upon a particular chromosome pair. Each human being carries two of these genes, which may be the same or different. Both A and B are dominant to O, which is recessive. A and B do not display dominance or recessiveness in regard to each other; if both are present both are expressed and the individual bearing them carries both agglutinogens in his red cells. The genetic constitutions corresponding to the four groups are as follows:

Group	Genetic constitution.
AB	AB
A	AA or AO
B	BB or BO
O	OO

From the point of view of anthropology the blood groups possess manifest advantages compared with any other measurements. In the first place the group to which a person belongs depends solely upon heredity, environmental influences having no effect upon the result. This is very different from cranial index, for example. Newman, Freeman and Holzinger (1937), in their studies on twins show clearly how markedly cranial measurements may be affected by non-genetic influences. In the second place, the mode of inheritance is simple. This is quite different from such a measurement as stature, which, though it is determined very largely by inheritance (Fisher, 1918; Soboleva and Ignatiev, 1936), is genetically complex, many genes being concerned. In the third place, the blood groups have at best a very small selective value. In fact, it has never yet been demonstrated that it is of any advantage or disadvantage to belong to a particular blood group. It is true that it is likely on theoretical grounds that slight differences do exist, Evidence for such a difference is presented in this thesis; but any such effect must be extremely small and of quite a different order of magnitude from the clear adaptive value of pigmentation, for example. Skin and hair colour could be radically changed by selection in relatively few generations; and as Haldane (1940) has pointed out, though studies of pigmentation may be valuable in elucidating recent changes, they cannot, unlike the blood groups, throw light upon the more remote origins of the races of mankind.

Neither A nor B is new, for both occur in the anthropoid apes. Theories which seek to account for existing differences in blood group frequencies by single mutations cannot explain the facts. Furthermore, mutation alone, even when repeated, cannot provide an explanation; Haldane has shown, for example, that a population devoid of B at the end of the last

Ice Age could not have acquired by mutation (at any rate it is reasonable to assume) as much B as is now present in the peoples of Western Europe.

The valuable collection of data by Boyd (1939) displays a clear pattern in the frequency of B. This is highest in Central Asia, in India and in China, and falls off in all directions from this centre. It is possible that the aboriginal populations of north America and Australia were devoid of B and that such few B genes as are now found are to be attributed to recent crossing. In two or three south American Indian tribes a very high frequency of B has been reported. The numbers tested are small, but it is quite possible that independent centres of high B do exist in that Continent. In general, however, it is difficult not to agree with Haldane (1940) that the great majority of B genes now in existence are to be attributed to a spread outwards from Central Asia.

The variations in frequency of A are just as strongly marked, though much more complex, and no broad general pattern is to be discerned. It is perhaps most probable that existing differences are to be referred principally to a time when man was a relatively rare animal, like the gorilla of to-day. At a time when a few scattered tribes, a few hundreds of individuals altogether, might represent the population of an area like the British Isles, it would be easy for the accidents of mutation and the chance association

??

of particular genes with characteristics of selective value to ensure that certain tribes were, for example, very high in A, while the individuals composing others were almost exclusively of group O. The subsequent rapid expansion of numbers, together with migration and crossing, would then account for the wide variations found to-day. That this view is correct is strongly supported by the fact that certain north American Indian tribes are extremely high in A while others occupying areas not far removed are equally high in O.

But though the origin of the blood groups is mysterious and theories to account for them cannot be advanced with any degree of assurance, it can be accepted for the reasons already given that studies on variations in frequency are of the greatest value in elucidating the relationship of existing populations, and are second to no other anthropological measurement in this respect.

The mass registration of blood donors in connexion with the war has provided data on a scale undreamed of hitherto. At an early stage, when these schemes were initiated, Prof. R.A. Fisher, Dr. G.L. Taylor, and their associates at the Galton Laboratory, University College, London, determined to make every effort to ensure that this unique opportunity should not be wasted, and that as far as was possible the data should be used for more general studies in addition to the primary purpose of providing blood and plasma for transfusion.

It should, perhaps, be pointed out here that few parts of the world offer greater possibilities for blood group studies than do these serologically variable islands; in few countries do such great variations occur within so small an area: variations which are related, moreover, to the different parent races which compose our present-day population.

The second opportunity provided by large bodies of data relates to the possibility of discovering whether there are not slight selective advantages associated with the blood groups. It is difficult to see how the existing polymorphism could have been perpetuated over such a period of time without postulating an equilibrium in which is concerned the balance of mutation rates and certain slight selective advantages and disadvantages. Hitherto, however, every attempt to relate membership of a particular blood group to any other characteristic, good or bad, has failed. It is clear that any differences must be very small and that relatively enormous numbers are needed for useful analysis. The present emergency has provided them for the first time. And it will be shown later that the data of the present paper do provide evidence pointing to a selective effect.

It has been my good fortune to have been in close touch with the serological work of the Galton Laboratory since its inception and to have had the opportunity of analysing, as my share in the co-operative effort, records relating to 120,000 donors from the six south-western counties of England. These, together with a sample from north Wales, of which some description has been given in two recent publications (Roberts, 1942a, 1942b), provide the material for the present thesis.

II. MATERIAL.

The west country records include 120,874 donors drawn from the six south-western counties of England: Gloucestershire, Wiltshire, Dorsetshire, Somersetshire, Devonshire and Cornwall, the area allotted to the Army Blood Transfusion Service.

The mass grouping of blood donors has on occasion led to records containing many wrong groupings. I am encouraged to believe, however, that this material compares favourably with other large bodies of data and that the proportion of wrong groupings is small. In the first place, the relative proportions of the four groups are closely similar to those found by the workers of the Galton Laboratory for others parts of southern England (Taylor, Race and Fisher, 1941). In the second place Bernstein's (1930) test yields a χ^2 which does not attain the level of significance: a favourable indication, and one sufficiently unusual for so large a number of donors. In the third place there are evidences of internal consistency which will become manifest during the course of this thesis.

I can provide direct figures in one case only. The reasons for this should perhaps be explained. In the practical business of obtaining blood and plasma for transfusion the important distinction is that between group O and the rest. Whole blood is obtained from donors of group O, while the blood of the remainder is pooled and used for the preparation of plasma. Wrongly grouped donors of group O will be discovered when they are bled. In the Bristol area over a considerable period the percentage of persons wrongly assigned to group O on original registration was one and a half.

The most common mistakes in grouping occur in connexion with groups B and AB (Taylor and others, 1939, 1940). As will be shown later there is evidence of this in the west country figures. Fortunately, however, it is O and A which are variable in frequency, and therefore important from the point of view of this thesis. Accordingly, nearly all the calculations and comparisons are based upon groups O and A only, group frequencies being used instead of calculated gene frequencies, which depend upon all four groups.

Finally, practically all the comparisons made are comparisons within the material. In each case I have had discussions with the officers of the Army Blood Transfusion Unit as to whether there was any possibility that errors in grouping were not distributed at random amongst the classes being compared. In one or two cases where this did not seem to be certain the comparisons have been omitted.

The great majority of groupings in connexion with the sample from north Wales were carried out, under conditions which should render mistakes very infrequent indeed, at the Regional Transfusion Centre in the Department of Pathology of the University of Liverpool.

III. GEOGRAPHICAL AND RACIAL VARIATIONS IN BLOOD-GROUP FREQUENCIES IN THE WEST COUNTRY AND IN WALES.

1. Previous information on the variation in the relative frequencies of O and A in the British Isles.

The remarkably large variations in blood-group frequencies in the British Isles are confined to the relative proportions of genes O and A. The frequency of gene B is about 6 per cent, corresponding to 8 or 9 per cent of persons falling into group B and 3 or $3\frac{1}{2}$ per cent into group AB. There appears to be little if any systematic variation in the frequency of B, though, of course, where the absolute frequency is so low relatively enormous numbers would be required in order to reveal moderate fluctuations.

As regards groups O and A the largest and most recent published figures are those of Taylor, Race and Fisher (1941). They are based on 150,000 donors. The samples from northern England and Scotland are smaller than those for southern England, but are quite large enough to yield reliable figures. The results are as follows:

	Percentage frequencies.	
	O	A
Southern England	45	43
Northern England	48	40
Scotland	52	34

Haldane (1940) quotes figures based on 784 Irish donors. The percentages of groups O and A are 56 and 30 respectively. (A much larger collection of figures, not yet published, from northern Ireland confirms this very low proportion of A. E.W. Hart's data. M.D. thesis, University of Cambridge, 1942).

Not only are the British Isles very variable in regard to O and A, but Scotland and Ireland are remarkable for the very low proportion of A. They yield the lowest figures in Europe and are comparable in this respect to Iceland, for which the figures given by Haldane (1940) are 56 and 32 respectively. As Haldane points out, the population of Ireland of to-day is probably similar, serologically speaking, to that of western Europe generally of 3000 years ago.

Fisher and Taylor (1940) have given an example of the light which the blood-groups may throw on the origins of peoples. It is clear from historical records that Scotland and Scandinavia were closely connected in the past, and that Iceland was colonized by the Scandinavian peoples. Yet the proportion of A in Scotland is, with Ireland and Iceland, the lowest in Europe, while the Scandinavian peoples are highest in this respect. It is difficult to escape the conclusion that the modern Scandinavians are quite different from those of Viking times, who must have been replaced by successive waves of immigrants from central and eastern Europe.

An alternative method of expressing the relative variations in the proportion of O and A is to give, as a percentage, the ratio A:O+A. This simply means the percentage of donors of group A ignoring groups B and AB. Reasons have already been advanced for preferring a method of this type to the calculation of gene frequencies from all four groups. The ratios are: southern England 49; northern England 45; Scotland 40; Ireland 35; Iceland 36.

2. Blood group frequencies in the six south-western counties of England.

The total count has yielded the following result:

O = 53,744, A = 53,150, B = 10,126, AB = 3,854,
Total = 120,874.

The corresponding percentage frequencies are:

O = 44.46, A = 43.97, B = 8.38, AB = 3.19.

The ratio A:O+A is 49.7.

The relative proportions of O and A are thus closely similar to those found by Taylor, Race and Fisher (1941) for the more easterly parts of southern England.

Following Bernstein's (1930) method, the gene frequencies are calculated as follows:

Let the group frequencies be denoted by O, A, B and AB.

Let the first estimates of the corresponding gene frequencies be denoted by r', p', and q'.

Let the final estimates be denoted by r, p, q, then:

$$r' = \sqrt{O}$$
$$p' = 1-\sqrt{(O+B)}$$
$$q' = 1-\sqrt{(O+A)}$$

Improved estimates are secured by putting $D=1-(r'+p'+q')$ and calculating:

$$r = (1+\tfrac{1}{2}D)(r' + \tfrac{1}{2}D)$$
$$p = p'(1+\tfrac{1}{2}D)$$
$$q = q'(1+\tfrac{1}{2}D)$$

The percentage gene frequencies in the west country are: r (gene O) = 66.72, p (gene A) = 27.32, q (gene B) = 5.96. In order to apply Berstein's test the expected group frequencies are calculated as follows:

$$O = r^2$$
$$A = p^2+2pr$$
$$B = q^2+2qr$$
$$AB = 2pq.$$

There is one degree of freedom, $\chi^2 = 2.566$, so P lies between 0.2 and 0.1. There is thus no significant discrepancy between these figures and the relative proportions expected on the accepted hypothesis of the inheritance of the blood-groups.

For so large a body of data the agreement between expectation and observation is encouraging, and indeed remarkable. It may be accepted as one piece of evidence that the basic figures do not contain many errors. Nevertheless, it should not be concluded on the basis of Bernstein's test alone that a body of data is necessarily satisfactory. There is only a single degree of freedom and two errors in opposite directions might balance each other. It is probable that this has happened in the present material in regard to group AB. It will be noted in Table 2 that the proportion of donors of this group in the Bristol area is 3.6 per cent. This is too high a figure but is balanced by too low a proportion in the other areas.

The records for the six counties were tabulated separately for nine areas based upon the following towns: Cheltenham, Bristol, Salisbury, Dorchester, Bridgwater, Exeter, Plymouth, Barnstaple and Truro. The group frequencies in each area are shown in Table 1. Table 2 shows the corresponding percentages together with the ratio A:O+A.

Table 1. Frequencies of the four blood groups in nine areas comprising the six south-western counties.

Area	O	A	B	AB	Total
Cheltenham	8,809	8,651	1,579	585	19,624
Bristol	17,939	17,984	3,347	1,470	40,740
Salisbury	2,831	2,921	551	196	6,499
Dorchester	1,851	1,755	291	103	4,000
Bridgwater	6,884	6,684	1,251	431	15,250
Exeter	8,425	8,325	1,726	594	19,070
Plymouth	390	372	76	18	856
Barnstaple	3,613	3,584	730	278	8,205
Truro	3,002	2,874	575	179	6,630
Total	53,744	53,150	10,126	3,854	120,874

Table 2. Percentage frequencies and ratio A:O+A in nine areas comprising the six south-western counties.

Area	O	A	B	AB	A:O+A
Cheltenham	44.9	44.1	8.0	3.0	49.5
Bristol	44.0	44.1	8.2	3.6	50.1
Salisbury	43.6	44.9	8.5	3.0	50.8
Dorchester	46.3	43.9	7.3	2.6	48.7
Bridgwater	45.1	43.8	8.2	2.8	49.3
Exeter	44.2	43.7	9.1	3.1	49.7
Plymouth	45.6	43.5	8.9	2.1	48.8
Barnstaple	44.0	43.7	8.9	3.4	49.8
Truro	45.3	43.3	8.7	2.7	48.9
Total	44.5	44.0	8.4	3.2	49.7

Much the most important question is whether there is any significant variation throughout the area in the relative proportions of groups O and A. Taking these two groups only and making a 2 x 9 comparison for homogeniety, $\chi^2 = 9.024$. There are 8 degrees of freedom, so P lies between 0.5 and 0.3. There is thus no evidence of any significant heterogeniety in the proportion of O and A throughout the six counties. For this reason it is unnecessary to give a map showing the boundaries of the nine sub-areas.

The proportions of groups B and AB seem to be distinctly variable. In the case of group B, comparing these donors with the remainder in each area, the 2 x 9 comparison yields a χ^2 of 26.404, which with 8 degrees of freedom gives P 0.001. Similarly, the 2 x 9 comparison in the case of group AB and the remainder gives a χ^2 of 47.771, an even more significant departure from homogeniety. Actually, however, the heterogeniety in the case of group AB is almost entirely due to the difference between the Bristol area and the rest. A 2 x 8 comparison, omitting the Bristol area, gives a χ^2 of 13.566, which does not attain the level of significance, P lying between 0.1 and 0.05. The heterogeniety in the case of groups B and AB is undoubtedly to be referred to errors in grouping and not to any real variation; it is a heterogeniety which is to be expected in almost any large sample. The errors may not be gross, but they do emphasize the desirability of confirming all the important comparisons to groups O and A.

3. Blood group frequencies in north Wales.

The north Welsh sample consists of 2,550 records of donors drawn from 17 centres situated in Caernarvonshire, Denbighshire and Flintshire. The total figures are as follows:

O = 1,224, A = 1,009, B = 234, AB = 83.

The corresponding percentage frequencies are:

O = 48.0, A = 39.6, B = 9.2, AB = 3.3.

The ratio A:O+A = 45.2

The calculated gene frequencies, using Bernstein's method, as described above, are:

O = 69.2, A = 24.4, B = 6.4.

Bernstein's test gives $\chi^2 = 0.183$. P lies between 0.7 and 0.5, so once again there is no significant departure from the expected proportions.

The proportions of O and A are closely similar to those found in northern England by Taylor, Race and Fisher (1941). An important separation of the material is, however, presented in a later section.

In order to examine homogeniety within the three counties in regard to O and A, it is necessary to anticipate and give the figures separately for donors bearing Welsh and non-Welsh family names: a distinction shown below to be of great importance. Table 3 shows the numbers falling into groups O and A at the 17 centres.

Table 3. <u>Number of donors of groups O and A at seventeen centres, distinguishing between Welsh and non-Welsh family names.</u>

	Welsh family names		Non-Welsh family names	
	O	A	O	A
Caernarvonshire:				
Caernarvon	42	27	17	13
Portmadoc	80	56	42	47
Bangor	67	53	85	81
Llandudno	41	19	152	125
Denbighshire:				
Colwyn Bay	28	14	47	62
Denbigh	22	17	12	11
Llangwyfan (Sanatorium)	15	5	6	8
Llanbedr (Sanatorium)	3	1	6	6
Ruthin	39	31	23	24
Wrexham	75	57	79	48
Flintshire:				
Rhyl	26	24	53	55
St Asaph	39	16	15	16
Flint	26	11	19	21
Mold	7	8	16	9
Holywell	3	-	7	5
A works	38	31	36	40
A factory	29	34	29	34
Total	580	404	644	605

In the case of the donors with non-Welsh family names a 2 x 16 comparison (adding together the figures for Llangwyfan and Llanbedr) gives $\chi^2 = 15.261$. There are 15 degrees of freedom, so P lies between 0.5 and 0.3. Agreement is excellent, so that within this area there is no evidence of heterogeneity amongst the non-Welsh donors.

In the case of the donors with Welsh family names (adding together the figures for Llangwyfan and Llanbedr, and also those for Mold and Holywell), $\chi^2 = 18.078$, so that for 14 degrees of freedom P is close to 0.2. If, however, the married women are excluded from the comparison, χ^2 becomes 24.048, which with 14 degrees of freedom gives P = 0.05 - 0.02. There is thus some evidence of heterogeniety. It is difficult, however, to discern any geographical pattern in the fluctuations. Perhaps the most important point is that variations in the proportions of O and A throughout the area amongst donors with Welsh names are trivial compared with the very large differences between these donors on the one hand and those with non-Welsh names on the other.

4. The use of family names in the further analysis of racial variations.

It was first suggested by Fisher and Vaughan (1939) that recent population movements can lead to an association between blood-group frequencies and family names. Out of 11,377 donors resident at Slough, they found that 591 with characteristically Welsh names were significantly lower in A than the remainder, the ratio A:O+A being respectively 45.3 and 49.5.

In the task of selecting family names I have found H.B. Guppy's "The Homes of Family Names" (1890) of the greatest value. This admirable book may, perhaps, not be as widely known as it deserves, though Bramwell (1923-4) relied upon it and pointed out its merits in connexion with some interesting observations on racial differences. Guppy's plan was simple. He selected the farming community as one which is relatively stable, and using Kelly's Post Office Directories determined county by county the frequency per 10,000 of all the surnames occurring. The Welsh counties are for the most part small, so he grouped them into two areas, north and south. In a few cases the smaller English counties were similarly grouped. Scotland was treated more briefly in an appendix. The book contains much additional information and analyses of the material, but its most valuable feature is an alphabetical index of 8000 family names giving the frequency per 10,000 of each name in each county. It should be added that Guppy ignored frequencies of less than 6 per 10,000.

5. The family name method: donors with Scottish and Irish names in the West country.

In order to test the family name technique in the present west country material the simplest plan is to make a selection of Scottish and Irish names to see whether their possessors are lower in A than the rest of the population. With so large a collection of data donors whose names have the prefixes Mac (including Mc and M') or O' provide a sufficiently large sample. The result is shown in Table 4; it is of a striking character.

Table 4. Table showing blood-group frequencies amongst donors resident in the west country bearing family names with prefix Mac or O'

	Frequencies				
	O	A	B	AB	Total
Names with prefix 'Mac'	421	306	65	31	823
Names with prefix 'O'	94	65	13	7	179
Total - six counties	53,744	53,150	10,126	3,854	120,874

	Percentage frequencies & ratio A:O+A				
	O	A	B	AB	A:O+A
Names with prefix 'Mac'	51.2	37.2	7.9	3.8	42.1
Names with prefix 'O'	52.5	36.3	7.3	3.9	40.9
Total - six counties	44.5	44.0	8.4	3.2	49.7

Donors bearing Scottish or Irish names are far lower in A than the population in which they are living. The figures are in fact similar to those found in Scotland and Ireland. It has to be remembered, however, that the figures for Scotland and Ireland would doubtless indicate a still lower proportion of A if donors with English names were excluded.

The difference between the Mac's and the remainder is, of course, highly significant. Making a 2 x 2* comparison for groups O and A only:

	O	A	Total
Mac's	421	306	727
Remainder	53,323	52,844	106,167
Total	53,744	53,150	106,894

$\chi^2 = 16.746$, there is one degree of freedom, and P is less than 0.001.

The difference is significant even in the case of the small sample of O's, for $\chi^2 = 4.632$ and P lies between 0.05 and 0.02.

This striking result shows that recent population movements can lead to a highly significant association between surnames and blood-group frequencies. Further examples are given later.

It is not true, of course, that race and family name are associated in every case. Some bearers will

* Yates' correction for continuity has been applied to all 2 x 2 tables in this thesis.

be of the indicated nationality; others partly so; yet others, for example some married women or descendants of immigrants very long resident in the area, will have none, or practically none, of the blood their surnames would appear to indicate. Nevertheless, if population movements are comparatively recent, a group of such donors will in the aggregate possess quite sufficient blood of the indicated race to reveal a difference, if it exists. The strongly positive results, such as that just given, together with the highly significant difference to be described later between the bearers of Welsh and English names resident in north Wales, make it legitimate to accept negative evidence also. In this thesis negative results are found in the case of donors with Welsh surnames resident in the Bristol and Cheltenham areas and also donors with Cornish names in Cornwall.

If there is any doubt as to whether the immigration in question in recent or not the sex-ratio will settle the point, for migrants are more frequently men than women. The Welsh donors, for example, in the Bristol and Cheltenham areas include a considerably higher proportion of men than the remainder of the donors in those areas.

is

but only if the sample taken is truly representative of the actual sex ratio of the region

The family name technique is as useful as it is simple. And its usefulness is by no means confined to blood-group studies. In any collection of material names are usually available, or can easily be made so. This simple separation will often provide results comparable to those which would be provided by infinitely more troublesome anthropological measurements.

6. Blood-group frequencies amongst the Welsh peoples.

With the help of Guppy's book (1890) the selection of Welsh family names in the north Welsh material has not proved a very difficult task, though a decision has not been easy in all cases. It has to be borne in mind in making the selection that Monmouthshire is as Welsh as the rest of Wales in regard to surnames, and Shropshire and Herefordshire hardly less so. Cheshire and Gloucestershire, on the other hand, do not show a high proportion of Welsh names. The family names range from those entirely confined to Wales in Guppy's list, through common ones certain to yield by migration a moderate frequency in some English counties, to those which though commoner in the Welsh areas than anywhere else, do have an independent origin and an appreciable frequency in various English counties. There are yet more extreme cases in which names though definitely Welsh are commoner in some parts of England than in Wales. In general most of those names which are both English and Welsh have been included. As the donors are resident in Wales the names must indicate a Welsh origin in the great majority of cases. Some of the names selected would, however, have to be excluded if Welsh names were being chosen in an English area.

A few examples may be mentioned to illustrate the more difficult cases. The name Harris is very widely spread throughout England and is relatively common in a number of counties. It is, however, considerably commoner in south Wales and Monmouthshire than anywhere in England; it also occurs in north Wales. The names James and Morris display much the same distribution but are more characteristically Welsh. Ellis is a north Welsh name, though it is as common, or commoner, in many English counties. Richards and Rogers are examples of names which have a wide distribution in England and are commoner in Cornwall than in Wales. Peters is a rare name listed by Guppy as occurring in north Wales, but also, with slightly greater frequency, in Cornwall and Somerset. All these have been included as Welsh names in Table 5. Certain names have been excluded because they are so much more often English than Welsh. Examples are Daniel, Rowland, Perkins, and Stephens; Stephens is three times commoner in Cornwall than in south Wales. Breeze, or Breese, provides a good example of a name having a double origin. In north Wales it is doubtless a contraction of ap Rees. But it also occurs in Norfolk and Suffolk, and there almost certainly has a different origin. It has been included.

Any difficulties in selection are, however, confined for the most part to names contributing very few donors. Twelve names only contribute 83 per cent of the total, and about them as well as a number of others there can be no doubt at all.

Table 5.

	No. of donors	Frequency per 10,000 (Guppy, 1890)		Classification.
		North Wales	South Wales	
Jones	256	1500	650	N . .
Williams	153	700	650	. E .
Roberts	112	500	110	N . .
Hughes	80	350	76	N . .
Davies	72	500	600	. E .
Owen	59	380	115	N . .
Evans	52	500	520	. E .
Edwards	36	150	140	. E .
Parry	31	66	22	N . .
Thomas	31	200	700	. . S
Lloyd	28	100	93	. E .
Griffiths	26	290	220	. E .
Griffith	14	Included with Griffiths, north Welsh form		N . .
Price	14	70	150	. . S
Ellis	13	25	-	N . .
Morris	11	80	76	. E .
Richards	11	70	93	. E .
Rowlands	10	40	27	. E .
Lewis	9	150	330	. . S
Pritchard	7	60	-	N . .
Rogers	7	18	32	. E .
Bellis	6	12	-	N . .
Elias	6	Not included		N . .
James	6	25	185	. . S
Pierce	6	30	-	N . .
Charles	5	Not included		N . .
Humphreys	5	75	-	N . .
Morgan	5	110	380	. . S
Phillips	5	30	150	. . S
Jenkins	4	30	220	. . S
Peters	4	9	-	N . .
Pugh	4	160	22	N . .
Cadwaladr	3	Not included		N . .
Francis	3	-	44	. . S
Harris	3	20	120	. . S
Humphries	3	Included with Humphreys		N . .

Table 5 (continued)

Powell	3	20	95	. . S
Rees	3	50	330	. . S
Breeze	2	50	-	N . .
Ffoulkes	2	Included with Foulkes		N . .
George	2	-	34	. . S
Probert	2	-	22	. . S
Prytherch	2	Not included		N . .
Vaughan	2	55	11	N . .
Bevan	1	-	55	. . S
Foulkes	1	25	-	N . .
Harries	1	Included with Harris		. . S
Hopkins	1	-	108	. . S
Howell	1	15	66	. . S
Howells	1	-	44	. . S
Hywel	1	Not included		. . S
Meredith	1	15	17	. E .
Morgans	1	Included with Morgan, north Welsh form		N . .
Tudor	1	40	-	N . .
Walters	1	-	60	. . S
Watkins	1	18	98	. . S
Wynn	1	Not included		N . .
Wynne	1	30	-	N . .
Total	1132			

Table 5 gives a list of the names selected as Welsh together with the number of donors contributed by each. The frequencies per 10,000 according to Guppy, in north and south Wales respectively, are also shown. The final column is based on these figures, the names being divided into those which are commoner in north Wales, those commoner in south Wales, and those which occur fairly equally in both parts of the country. The names Elias, Charles, Cadwaladr and Prytherch are not mentioned by Guppy. The names Hywel and Wynn are obvious variants of names usually differently spelt.

Table 6 shows the frequencies for donors with Welsh and with non-Welsh family names, distinguishing between men, single women and married women. In the great majority of cases donors were clearly distinguished as 'Mr', 'Mrs' and 'Miss'. Titles did cause ambiguity in a few cases, however. Women distinguished by 'Sister' or 'Nurse' have been included with the single women. The lists showed initials and family name only in a few instances; these donors have been included with the men. In the case of a few women Christian name and surname only had been entered without prefix; these have been included with the single women. The number of cases possibly misclassified is very small.

Table 6. Blood group frequencies amongst donors with Welsh and non-Welsh family names.

	O	A	B	AB	Total
1. Welsh family names:					
Men	270	190	51	10	521
Single women	209	128	37	14	388
Married women	101	86	27	9	223
Total	580	404	115	33	1,132
2. Non-Welsh family names:					
Men	238	215	51	12	516
Single women	270	243	39	23	575
Married women	136	147	29	15	327
Total	644	605	119	50	1,418

Table 7. Percentage frequencies of the blood groups amongst donors with Welsh and non-Welsh family names. Also the ratio A:O+A as a percentage.

	Percentage frequencies				A:O+A
	O	A	B	AB	
1. Welsh family names:					
Men	51.8	36.5	9.8	1.9	41.3
Single women	53.9	33.0	9.5	3.6	38.0
Men & single women	52.7	35.0	9.7	2.6	39.9
Married women	45.3	38.6	12.1	4.0	46.0
Total	51.2	35.7	10.2	2.9	41.1
2. Non-Welsh family names:					
Men	46.1	41.7	9.9	2.3	47.5
Single women	47.0	42.3	6.8	4.0	47.4
Married women	41.6	45.0	8.9	4.6	51.9
Total	45.4	42.7	8.4	3.5	48.4

Table 6 shows the counts for donors with Welsh and non-Welsh family names distinguishing between men, single women and married women. Table 7 shows the results as percentages and also shows the ratio A:O+A expressed as a percentage. The difference between the Welsh and non-Welsh donors is very striking. Taking the simplest comparison, a 2 x 2 table* showing groups O and A only, χ^2 is 11.813, which for one degree of freedom gives P = 0.001 - 0.0001, so that the

* Yates' correction for continuity has been used, as in all the other fourfold tables.

difference is highly significant.

In the case of the non-Welsh names a 3 x 2 table showing numbers of O and A for men, single women and married women respectively, yields a χ^2 of 1.800. There are two degrees of freedom and P lies between 0.5 and 0.3. There is no evidence of heterogeniety, therefore, in the relative proportions of O and A amongst the three classes of donors.

The men and single women with Welsh names are very similar in the percentage of O and A. In the fourfold table $\chi^2 = 0.763$ and P lies between 0.5 and 0.3, so there is no evidence of heterogeniety. The married women with Welsh family names are, however, distinctly higher in A and lower in O, as would of course be expected if some of them are English women married to Welsh husbands. The fourfold table giving numbers of O and A for married women and for the remainder gives $\chi^2 = 2.076$, P lying between 0.2 and 0.1. As, however, the deviation is in the expected direction, the difference verges on significance.

The ratio A:O+A for the non-Welsh donors is 48.4, a figure closely similar to that found by Taylor, Race and Fisher (1941) in southern England. The figure of 41.1 for donors with Welsh family names (or 39.9 if married women are excluded) is similar to the 40 found by these observers in Scotland. Presumably, however, the figure for Scotland would be still lower if donors with English names were eliminated.

The ratio AB:B+AB confirms the difference between donors with Welsh and English family names, the figure being distinctly lower amongst the former. The donors with English names show 29.6 per cent and those with Welsh names 22.3 per cent.

At first sight it may seem surprising that so high a proportion of the north Welsh donors, no less than 56 per cent, should bear English family names. This is largely explained by the choice of centres. The north Welsh coast has become in recent times a Lancashire suburb and dormitory. Llandudno, Colwyn Bay and Rhyl can now hardly be described as Welsh towns. There has also been, of course, a considerable movement of population into north Wales as the result of the war. In the present material Llandudno makes the largest contribution of any centre and only 18 per cent of these 379 donors bear Welsh names. In this area of Britain two populations live side by side, still widely separated by language and tradition. It is in just such an area that the family name method is likely to reveal the largest differences.

In Table 8 the Welsh names are divided into three classes, as previously shown in Table 5, which gives the frequency per 10,000 (according to Guppy) in both areas. Names which are more than twice as common in the north as in the south are classified as north Welsh names; those twice as common in the south as south Welsh; where the difference is less than this

Table 8. Further subdivision of donors with Welsh family names.

1. Numbers of donors:	O	A	B	AB	Total
North Welsh names	356	201	48	22	627
Names of equal frequency	186	161	52	8	407
South Welsh names	38	42	15	3	98
2. Percentage frequencies and ratio A:O+A	**O**	**A**	**B**	**AB**	**A:O+A**
North Welsh names	56.8	32.1	7.7	3.5	36.1
Names of equal frequency	45.7	39.6	12.8	2.0	46.4
South Welsh names	38.8	42.9	15.3	3.1	52.5

the names are classified as being of approximately equal frequency. The heterogeneity is highly significant. Comparing O and A for the three types of family name, χ^2 is 14.105, P for two degrees of freedom lying close to 0.001. The magnitude of the difference is somewhat startling. On such numbers as these, however, the fiducial limits are wide and some of the effect may be due to chance. But there can be no doubt that there is a significant difference and that donors bearing names characteristic of south Wales are higher in A than those whose names are more frequent in the north.

This result at once raises the question as to whether or not the population of south Wales resembles that of north Wales, with Scotland and Ireland, in being low in A. The Bristol and Cheltenham figures from the west country data contain many Welsh names and it is certain that owing to proximity to south Wales, as well as owing to the much greater population of that part of the country, these names must be derived, recently or remotely, from south Wales. And this will be so even in the case of names commoner in north Wales than in south Wales.

In choosing Welsh names in an English area certain of them can no longer be included, for while it is legitimate to assume that in Wales a name which is definitely Welsh but which also occurs in some English counties will have a Welsh origin in the great majority of cases, this does not follow in an

English area. Accordingly the following criteria have been adopted. List 1 is composed of names shown by Guppy to be at least 5 times as common in either north or south Wales as in any English county (in the farming communities of the 'eighties); List 2 contains a further selection of names which are at least twice as common. Frequencies in Shropshire and Herefordshire have been ignored in preparing both lists.

The names included, with the number of donors contributed by each, are as follows:

List 1. Anthony (7), Bevan (47), Bowen (16), Davies (218), Evans (259), Griffith (2), Griffiths (112), Harry (5), Howells (14), Hughes (120), Jenkins (154), John (10), Jones (603), Llewellin (14), Llewellyn (25), Lloyd (59), Matthias (1), Morgan (227), Morgans (3), Owen (71), Owens (13), Preece (12), Probert (9), Pugh (19), Rees (51), Roberts (159), Rowlands (11), Ryder (7), Tudor (9), Vaughan (32) and Wynne (4).

Total = 2,293.

List 2. Breeze (1), Edwards (181), Hopkins (52), Howell (43), Humphreys (8), Humphries (49), James (233), Lewis (294), Parry (32), Phillips (176), Pierce (5), Powell (168), Price (168), Pryce (7), Thomas (323), Walters (54), Watkins (84), and Williams (491). Total = 2,369.

The results are shown in Table 9.

Table 9. Comparison with the remainder of donors with Welsh names resident in the Bristol and Cheltenham areas.

	Frequencies.				
	O	A	B	AB	Total
Bristol area:					
Total	17,939	17,984	3,347	1,470	40,740
Welsh names					
List 1	688	700	150	69	1,607
List 2	719	727	127	62	1,635
Cheltenham area:					
Total	8,809	8,651	1,579	585	19,624
Welsh names					
List 1	287	301	73	25	686
List 2	324	325	64	21	734

	Percentage frequencies & ratio A:O+A				
	O	A	B	AB	A:O+A
Bristol area:					
Total	44.0	44.1	8.2	3.6	50.1
Welsh names					
List 1	42.8	43.6	9.3	4.3	50.4
List 2	44.0	44.5	7.8	3.8	50.3
Cheltenham area:					
Total	44.9	44.1	8.0	3.0	49.5
Welsh names					
List 1	41.8	43.9	10.6	3.6	51.2
List 2	44.1	44.3	8.7	2.9	50.1

The donors with Welsh family names are actually slightly higher in A than the population amongst which they live, though the differences are not, of course, significant. Yet, as was shown in Table 4, this same material has yielded a large and significant difference in the case of Scottish and Irish names. The numbers of Welsh donors are large enough to reveal a very moderate difference - if it existed.

It might be argued that a high proportion of persons with Welsh names living in the Bristol and Cheltenham areas belong to families long resident in those localities, so that the original contribution of many O genes by their ancestors has become spread throughout the population through intermarriage. The sex-ratio, however, which reveals a large and significant difference in both areas shows that much of the immigration must be very recent.

	Males per 100 females.
Bristol area:	
Donors with Welsh names	84.1
Remainder	74.3
Cheltenham area:*	
Donors with Welsh names	147.8
Remainder	122.2

I am indebted to Prof. Fisher for some hitherto unpublished figures from south Wales. A sample of donors from Cardiff gives the following result:

* The higher proportion of men in this area is due to the inclusion of a large number of employées of the Great Western Railway at Swindon.

O = 798, A = 750, B = 161, AB = 56.

The corresponding percentages are:

O = 45.2, A = 42.5, B = 9.1, AB = 3.2.

The ratio A:O+A is 48.4.

A sample from Swansea gives the following figures:

O = 240, A = 214, B = 61, AB = 22.

The percentages are:

O = 44.7, A = 39.9, B = 11.4, AB = 4.1.

The ratio A:O+A is 47.1.

Certain discrepancies between the sexes, however, in the case of Swansea make the use of these last figures somewhat questionable. These populations are very little lower in A than the population of southern England generally, for which the ratio is 49.

It might be suggested that the proportion of persons of non-Welsh ancestry in Cardiff and Swansea is now large enough to obscure the low contribution of A made by persons of Welsh blood. The family name technique would be very valuable in this connexion. I greatly doubt, however, whether the English influx into Cardiff and Swansea is nearly so great as into north Wales, particularly into the north Welsh coastal area, so strongly represented in the north Welsh sample. Yet Cardiff shows significantly more A than the north Welsh material taken as a whole.

The results for the Bristol and Cheltenham areas alone make it probable that any difference between the inhabitants of south Wales, even those bearing Welsh names, and the inhabitants of southern England must be at best quite small. This conclusion is confirmed by the figures provided by Prof. Fisher for Cardiff and Swansea and by the difference noted in the north Welsh material between donors with north and south Welsh names. Further evidence is of course desirable, and several additional points could usefully be investigated. In particular it would be interesting to determine the boundaries of the area of low A. But even on the evidence now available it seems highly probable that the inhabitants of north and south Wales, even those bearing the same names, differ very widely in blood-group frequencies.

This somewhat unexpected result may not after all be so very surprising. It is known that different influences have affected the two parts of the country. Furthermore, communications between north and south must always have been difficult, as indeed they are to-day. Some years ago I heard it said in this connexion that the best meeting place for all Wales was Shrewsbury in summer and London in winter. Surnames came into common use in Wales during the sixteenth century, so if the present findings are confirmed it must follow that the peoples who happened at that time to adopt the same family names were already very different. From the point of view of blood-groups then, and there is no more valuable anthropological guide, it is probably not going too far to conclude that while the north Welsh are kind to the Highland Scots and the Irish, the south kin

Welsh, either from the period of the Saxon invasions, or by subsequent infiltration prior to Tudor times, are almost one with the southern English.

7. Blood-group frequencies amongst the Cornish people.

Except for the inclusion of two or three small Cornish centres in the Plymouth and Barnstaple areas, the area centred on Truro includes the whole of the county. The figures have been given in Tables 1 and 2 and it has been shown that there is no significant heterogeneity in the proportions of O and A throughout the six counties, including Cornwall. If a single comparison is made the figures are as follows:

	O	A
Truro area:	3,002	2,874
Remainder of six counties:	50,742	50,276

This fourfold table yields a χ^2 of 1.603. There is one degree of freedom and P lies between 0.3 and 0.2, so there is no evidence that Cornwall is any lower in A than the adjacent counties.

In order to apply the family name technique use was again made of Guppy's book. Unlike Wales, Cornwall shares many names with other English counties. Nevertheless, this county (together with Devonshire) is richer in names peculiar to itself than any other part of England. Guppy gives a long list of names which do not occur with appreciable frequency in any other county. The numbers of donors of group O and A amongst the bearers of these names and the remainder of the Truro donors are as follows:

	O	A	A:O+A
Bearers of names peculiar to Cornwall:	322	322	50.0
Remainder of Truro donors:	2,680	2,552	48.8

The result is unequivocal, for donors with names peculiar to Cornwall are actually slightly higher in A than the remainder. In view of the success of the family name technique in revealing associations where they do exist, as in the case of Scottish and Irish donors in this same material, as well as in the case of the north Welsh sample, it can be concluded that the people of Cornwall of the present day do not differ from their neighbours in the frequency of A and bear no relationship in this respect to the Scots, the Irish and the north Welsh.

IV. A SEARCH FOR EVIDENCE AS TO WHETHER THE BLOOD-GROUPS POSSESS SELECTIVE VALUE.

1. General.

It was pointed out in the introduction that in spite of many enquiries no association has yet been demonstrated between membership of a particular group and any other characteristic, good or bad. Yet it is difficult to think that the blood-groups have no selective value. A polymorphism which probably dates from pre-human times, for it is found to-day in the anthropoid apes, is unlikely to have been perpetuated in the absence of a mechanism involving differences in reproductive rates.

As a simple example of the type of mechanism which could account for the diversity observed, it might be supposed that it was a slight advantage to be a heterozygote. This hypothesis can be tested, for while all members of group O are homozygous, bearing two O genes, the majority of persons of group A are heterozygous, about 5 out of 6 of them carrying gene A on one chromosome of the pair concerned and gene O upon the other. In view of previous failures it is clear, however, that any selective effect is at best a very small one; this would in any case be anticipated on general grounds. Relatively very large numbers indeed are needed and while numbers even of the order secured in connexion with transfusion schemes might still not be adequate, at least they provide material on a scale not approached hitherto.

The large sample of 120,000 donors from the west country provides the opportunity for exploring two possibilities: first, that there is a connexion between the blood-groups and age; secondly, that there is a sex-difference.

2. Blood-groups and age.

It was intended that the ages of donors should be entered on their record cards. This had been done in the great majority of cases in most areas. In the case of the Cheltenham area, however, so many were omitted* that it seemed not worth while tabulating the ages of the remainder. In the case of the Truro area it was decided to await a further drive for donors in view of the importance of securing an adequate sample in order to see whether Cornwall differed from the rest of south-western England in respect of O and A. By this time a new system of registration was in operation, ages being no longer recorded. The sample of 6,600 finally obtained replaced a previous sample of about 1,000, tabulated at the same time as the rest of the material. The records of the 1,000 could not be used because some of the donors were the same as those included in the larger group of 6,600. Their ages had been recorded, however, so they are available for the purposes of the present section. Altogether ages were available in the case of 55,338 women and 30,100 men.

Age distributions are given in Table 10.

* At the Great Western Railway's works at Swindon.

Table 10. Age distributions of donors.

Age	Women					Men				
	O	A	B	AB	Total	O	A	B	AB	Total
10-19	3,517	3,429	640	259	7,845	1,332	1,330	253	101	3,016
20-29	7,240	6,960	1,296	512	16,008	3,040	3,079	559	216	6,894
30-39	5,676	5,815	1,159	457	13,107	4,135	4,048	729	317	9,229
40-49	4,368	4,239	799	279	9,685	2,471	2,529	509	205	5,714
50-59	2,807	2,789	547	231	6,374	1,628	1,531	300	116	3,575
60-69	940	933	184	80	2,137	688	652	109	43	1,492
70-79	73	81	16	3	173	67	90	6	7	170
80-89	4	4	1	-	9	3	5	1	-	9
90-99	-	-	-	-	-	1	-	-	-	1
Total	24,625	24,250	4,642	1,821	55,338	13,365	13,264	2,466	1,005	30,100

The mean ages, with standard errors, of women and men of the four groups are shown in Table 11.

Table 11. Mean ages with standard errors of donors of the four blood-groups.

Group	Women		Men	
	Mean age.	Standard error.	Mean age.	Standard error.
O	34.529	0.088	36.778	0.117
A	34.646	0.088	36.653	0.117
B	34.866	0.202	36.622	0.266
AB	34.797	0.326	36.751	0.417
Total	34.617	0.058	36.709	0.078

The differences are very small in both sexes. That there is no significant association between mean age and blood-group is most conveniently demonstrated in the form of an analysis of variance. This is shown in Table 12.

Table 12. Analysis of Variance - Mean ages and blood-groups.

	Women		
	Degrees of freedom	Sum of squares	Mean square
Between blood-groups	3	559.3	186.4
Within blood-groups	55,334	10,470,841.8	189.2
Total	55,337	10,471,401.1	-

	Men		
	Degrees of freedom	Sum of squares	Mean square
Between blood-groups	3	125.0	41.7
Within blood-groups	30,096	5,450,031.7	181.1
Total	30,099	5,450,156.7	-

In the case of both sexes the mean square within groups exceeds the mean square between groups, so there is no evidence of any significant age differences between donors of the four groups. This means that over the age range covered, essentially 18-60, there is no evidence that donors of any particular group survive on the average for a longer or shorter period than those of any other. At these ages, therefore, if any selective difference does exist it must be of an extremely small order of magnitude. Naturally, the figures yield no information as to whether there might not be differences before birth and during infancy and childhood, during the first part of which period the death rate is high. Nor do they cover adequately the period of old age, though this is less important, as differences at that end of the life cycle have little effect on reproductive rates.

2. Blood-groups and sex.

Here for the first time a difference emerges. In the west-country material women tend to be more often of group O, men of group A. The results for the whole area are as follows:

	O	A
Women	33,042	32.225
Men	20,702	20,925

In terms of percentages (ignoring groups B and AB), 49.37 per cent of women fall into group A, compared with 50.27 per cent of men.

The difference is, of course, a small one; but it is amply significant on such relatively large numbers. The 2 x 2 table above yields a χ^2 of 8.086, P lying between 0.01 and 0.001.

It is very important to examine the homogeniety of the difference throughout the area, for such factors as migration might lead to a spurious association. Homogeniety can be tested by calculating χ^2 separately for each of the 9 areas. This yields 9 degrees of freedom. The χ^2 for the total is deducted. The remainder, which corresponds to 8 degrees of freedom, then represents heterogeniety in one rational sense. The data for making the calculation are given in Table 13.

Table 13. <u>Donors of groups O and A amongst women and men in the nine south-western sub-areas</u>.

Area	Women		Men	
	O	A	O	A
Cheltenham	4,049	3,775	4,760	4,876
Bristol	10,283	10,237	7,656	7,747
Salisbury	2,045	2,119	786	802
Dorchester	1,257	1,228	594	527
Bridgwater	4,392	4,197	2,492	1,487
Exeter	6,279	6,132	2,146	1,193
Plymouth	316	276	74	96
Barnstaple	2,577	2,510	1,036	1,074
Truro	1,844	1,751	1,158	1,123
Total	33,042	32,225	20,702	20,925

The analysis is as follows:

	Degrees of freedom	χ^2	P
Sum for nine areas	9	21.096	0.02-0.01
Total for whole area	1	8.086	0.01-0.001
Difference attributable to heterogeniety	8	13.010	0.2 -0.1

To complete the evidence each sex can be tested separately for heterogeniety in the relative proportions of O and A throughout the area. In the case of the women $\chi^2 = 13.284$. There are 8 degrees of freedom and P lies between 0.2 and 0.1. In the case of the men $\chi^2 = 9.488$ and P lies between 0.5 and 0.3.

The whole of the heterogeneity can be attributed to the difference between the sexes; there is no evidence of heterogeneity due to any other cause; nor of variation in this respect throughout the area. It can be concluded, therefore, that the figures for the six counties agree in indicating that women fall more often into Group O and less often into Group A than do men.

The reality of this sex difference is supported by records, not yet published, which Prof. R.A. Fisher has obtained from Yorkshire; there, too, precisely the same effect emerges.

It is almost impossible to conceive of a genetic mechanism of any plausibility which could account for the difference. It is much more likely to be due to a slightly lower survival value in the case of persons of group O. It is true that the analysis by age recorded in the previous section lends no support to this hypothesis, but the ages affected are for practical purposes 18-60 only. Pre-natal mortality is high and so is immediate post-natal mortality. It is known that the death rate amongst boys at these ages greatly exceeds that of girls (Crew 1937); in fact the sex-ratio at the time of conception may be no less than 140 boys per 100 girls. If during these critical stages O babies were more likely to die than A babies, then differential mortality as between the sexes would lead to the survival into childhood and adult life of a higher proportion of O girls as compared with boys, and so produce the result found in the west country material.

The underlying explanation could well be that it is a slight advantage to be a heterozygote. As was pointed out above this is just the kind of mechanism which could have preserved a polymorphism over so long a period. In the case of persons of group O all are homozygous, bearing twin O genes. The majority of persons of group A are, on the other hand, heterozygous; for each person with two A genes there will exist nearly five heterozygotes bearing one A gene and one O gene.

The hypothesis can be put to the test by grouping still-born children and those dying young, a task which would now present no difficulty to the serologist. In the case of miscarriages indirect evidence could be obtained by grouping the parents. Relatively moderate numbers should suffice to show whether this is indeed the solution of one of the mysteries associated with the blood-groups.

V. SUMMARY AND CONCLUSIONS.

1. The material analysed in this thesis consists of records of 120,874 blood donors drawn from the six south-western counties of England, together with a sample of 2,550 donors from three north Welsh counties.

2. The percentage frequencies of the blood-groups in the west country are:

O = 44.5, A = 44.0, B = 8.4, AB = 3.2.

The corresponding gene frequencies are:

O = 66.7, A = 27.3, B = 6.0.

Bernstein's test does not reveal any significant departure from expectation. There is no significant variation in the relative proportions of O and A throughout the area.

3. The percentage frequencies of the blood groups in the north Welsh counties are:

O = 48.0, A = 39.6, B = 9.2, AB = 3.3.

The gene frequencies are:

O = 69.2, A = 24.4, B = 6.4.

Bernstein's test does not reveal any significant departure from expectation. Amongst donors with English family names there is no evidence of heterogeneity in the proportion of O and A throughout the area. In the case of donors with Welsh family names there is some evidence of slight heterogeneity. This does not follow any simple geographical pattern and is trivial compared to the large difference between donors of the two classes.

4. The suggestion of Fisher and Vaughan (1939)

that recent population movements can lead to association between blood-group frequencies and family names is amply substantiated. This simple method of study is shown to possess great value in the eludication of racial differences. In the west country material donors bearing names with the prefixes "Mac" and "O" are much lower in A than the remainder.

5. In the north Welsh material the percentage of donors of groups O and A are:

	O	A
Welsh family names	51.2	35.7
Non-Welsh family names	45.4	42.7

The difference is highly significant. The former figures are similar to those found in Scotland and Ireland, the latter to the figures for southern England.

6. Donors with names more characteristic of north Wales are significantly lower in A than donors with names more characteristic of south Wales. Donors with Welsh names resident in the Bristol and Cheltenham areas (and doubtless overwhelmingly of south Welsh origin) are no lower in A than the remainder of those populations. Samples from Cardiff and Swansea do not reveal a low proportion of A. It is concluded that there is strong evidence for believing that while the north Welsh are kin, serologically speaking, to the Scots and the Irish, sharing with them the lowest proportion of A in Europe, the southern Welsh are almost indistinguishable from the southern English.

7. Cornish donors, even those bearing characteristically Cornish names, are no lower in A than the

rest of the population of southern England.

8. There are no significant age differences between donors of the four groups, either in the case of men or of women. This indicates that if any selective differences distinguish the blood-groups they must be minute between the ages of about 18 to 60 at least.

9. In the south-western material there is a sex-difference. Women, as compared with men, are slightly more often of group O and less often of group A. The tentative hypothesis is advanced that here, for the first time, is evidence that the blood-groups do possess selective value. The difference would be explained if at the early critical stages of high mortality, when considerably more boys are dying than girls, A children had a greater chance of survival than O children. This effect cannot persist into adult life, however, as is shown by the absence of any difference in mean ages between donors of the four groups. The sex-difference might be due to homozygotes being at a slight disadvantage compared with heterozygotes; this is precisely the kind of mechanism which would explain the persistence of a polymorphism as old as the species.

VI. ACKNOWLEDGEMENTS.

The west country records are those of the Army Blood Transfusion Service and I find it difficult to acknowledge adequately my debt to Brigadier L.E.H. Whitby, C.V.O.,M.C., and his officers, not only for making the material available but for providing admirable facilities and for so readily giving help and advice at all times. The analysis of such large numbers occupied several months; it was a big task, but this unfailing kindness made it a very pleasant one. In addition to Brigadier Whitby, I should like to mention particularly Major B. Lacey, Capt. T. Fortescue Fox and Mr. R.G. McCrudden. Once again I have to record my indebtedness to my assistant Mrs. Joan Lenton for her careful and accurate work.

I am greatly indebted to Dr. R.P. Picton Davies, of Ruthin Castle, north Wales, and Dr. J. Emyr Jones, Assistant Medical Officer of Health for Caernarvonshire, who very kindly provided the north Welsh material. The donors were grouped in connexion with the E.M.S. Blood Transfusion Service, the majority of the groupings having been performed at the Regional Transfusion Centre in the Department of Pathology of the University of Liverpool; I wish to express my thanks to Prof. T.B. Davie, the Regional Blood Transfusion Officer, for so readily allowing the data to be made available. I am also greatly indebted to Prof. R.A. Fisher for providing the hitherto unpublished figures for Cardiff and Swansea.

VII. REFERENCES.

F. Bernstein (1924) Ergebnisse einer biostatistischen zusammenfassenden Betrachtung uber die erlichen Blutstrukturen des Menschen. Klin. Woch. 3: 1495.

__________ (1925) Zusammenfassende Betrachtungen uber die erblichen Blutstruckturen des Menchen. Zeit. f. indukt. Abstamm.-u. Vererbungsl. 37: 237.

__________ (1930) Fortgesetzte Untersuchungen aus der Theorie der Blutgruppen. Zeit. f. indukt. Abstamm. u. Vererbungsl. 56: 233.

W.C. Boyd. (1939) Blood-groups. Tabulae Biologie. 17: 113.

B.S. Bramwell (1923-4) Observations on racial characteristics in England. Eugen. Rev. 15: 480 and 556.

F.A.E. Crew (1937) Brit. Assoc. for the Advancement of Science. Report of Annual Meeting. p. 95.

T. Dobzhansky (1941) Genetics and the Origin of Species. 2nd Edition. New York: Columbia University Press.

R.A. Fisher (1918) The correlation between relatives on the supposition of Mendelian inheritance. Trans. Roy. Soc. Edin. 52:

__________ (1941) Statistical Methods for Research Workers. 8th Edition. Edinburgh: Oliver & Boyd.

R.A. Fisher & G.L. Taylor (1940) Scandinavian influence in Scottish ethnology. Nature. 145: 590.

R.A. Fisher & J. Vaughan (1939) Surnames and blood-groups. Nature. 144: 1047.

H.B. Guppy (1890) The Homes of Family Names. London: Harrison & Son.

J.B.S. Haldane (1940) The blood-group frequencies of European peoples, and racial origins. Human Biology. 12: 457.

L. Hirszfeld & H. Hirszfeld (1919). Serological differences between the blood of different races. The results of researches on the Macedonian front. Lancet. 197 (2): 675.

H.H. Newman, F.N. Freeman & K.J. Holzinger (1937) Twins: A Study of Heredity and Environment. Chicago: University of Chicago Press.

J.A.F. Roberts (1942a) Surnames and blood-groups, with a note on a probable remarkable difference between north and south Wales. Nature. 149: 138.

____________ (1942b) Blood-group frequences in north Wales. Ann. Eugen. 11 : 260.

G.V. Soboleva & M.V. Ignatiev (1936) The roles of heredity and environment in variability of height and weight. Proc. Maxim Gorky Medico-Genetical Inst. Moscow. 4: 376.

G.L. Taylor & E.W. Ikin (1939) Observations on the performance of blood-group tests. Brit. Med. Jour. 1 : 1027.

G.L. Taylor, R.R. Race, A.M. Prior & E.W. Ikin (1940) The weak "A" reaction found in some cases of the blood-group "AB". Brit. Med. Jour. 1: 297.

G.L. Taylor, R.R. Race & R.A. Fisher (1941) Report of Meeting. Brit. Med. Jour. 1: 315.

LIST OF PUBLISHED WORKS SUBMITTED TOGETHER WITH MANUSCRIPT PART OF THESIS.

Book.

An Introduction to Medical Genetics. Oxford Medical Publications. 1940.

Papers.

Studies on a Child Population.

I. Definition of the sample, method of ascertainment and analysis of the results of a group intelligence test* (With R.M. Norman & Ruth Griffiths). Annals of Eugenics. Lond. 6: 319. 1935.

II. Retests on the Advanced Otis and Stanford-Binet Scales with notes on the use of a shortened Binet Scale (with Ruth Griffiths). Ibid. 8: 15. 1937.

III. Intelligence and Family Size. (with R.M. Norman & Ruth Griffiths). Ibid. 8: 175. 1937.

IV. The form of the lower end of the frequency distribution of Stanford-Binet intelligence quotients and the fall of low intelligence quotients with advancing age. (with R.M. Norman & Ruth Griffiths). Ibid. 8: 319. 1937.

V. The resemblance in intelligence between sibs. Ibid. 10: 293. 1940.

Sex-linked microphthalmia sometimes associated with mental deficiency. Brit. Med. Jour. 2: 1213. 1937.

A study in the variability in the incidence of rheumatic heart disease within the City of Bristol. (with C. Bruce Perry). Brit. Med. Jour. (Supplement) 2: 154. 1937.

The place of genetics in the practice of medicine. Newcastle Med. Jour. 17: 115. 1937.

Intelligence and family size. Eugen. Rev. 30: 237. 1939.

Observations on a representative group of children of school age, with an account of some family and social characteristics of the brightest, the average and the dullest. Proc. Amer. Assoc. on Mental Deficiency. 44: 79. 1939.

---

* Written before taking the degrees of M.B.,Ch.B., but included for the sake of completeness.

Does poliomyelitis affect intellectual capacity: an investigation on 98 cases. (with R.G. Gordon and Ruth Griffiths). Brit. Med. Jour. 2: 803. 1939.

Surnames, intelligence and fertility. Nature. 145: 939. 1940.

Inheritance of mental deficiency. Proc. Seventh International Genetical Congress. 249. 1941.

Surnames and blood-groups, with a note on a probable remarkable difference between north and south Wales. Nature. 149: 138. 1942.

Blood-group frequencies in north Wales. Ann. Eugen. London. 11: 260. 1942.